Investigating Ad Transparency Mechanisms in Social Media

Oana Goga
CNRS researcher at University of Grenoble Alpes
(work done with Athanasios Andreou, Ghiridhari Venkatadri, Patrick Loiseau, Krishna Gummadi, Alan Mislove)

16 Octobre 2017
Entretien Jacques Cartier
Why did I, a user, received this particular ad?
Responsible processes for seeing an ad

- The ad targeting process
- The data inference process
Ad targeting process

- Motherhood
- New mover
  - Likely to engage in Politics (Liberal)

Advertiser

+ bid amount
optimizations

Ad platform

available users

(U1, Ad1)
(U2, Ad2)
(U3, Ad3)

other advertisers
Ad targeting process

The reasons why someone received a particular ad are the result of a complex process!
More challenges in constructing explanations

• Many design choices:
  ▶ format, length, amount of details, …

• What is a good explanation?
  ▶ promote trust
  ▶ satisfy curiosity
  ▶ detect malicious or deceiving targeting
  ▶ verify compliance
Explanations of the ad targeting process
Explanations of the ad targeting process

Why Am I Seeing This Ad?

One reason you're seeing this ad is that STAEDTLER wants to reach people interested in Etsy, based on activity such as liking Pages or clicking on ads.

There may be other reasons you're seeing this ad, including that STAEDTLER wants to reach people ages 18 and older who live in Germany. This is information based on your Facebook profile and where you've connected to the internet.
Desired properties of an ad explanation

audience = (Montreal & Canada) or EJC or not France

U₁ = {Montreal, Canada, Fishing, Driving}
U₂ = {EJC, Hiking, Skiing}
U₃ = {not France, Cuisine}
U₄ = {Montreal, Canada, Reading}
Desired properties of an ad explanation

audience = (Montreal & Canada) or EJC or not France

U1 = {Montreal, Canada, Fishing, Driving}
U2 = {EJC, Hiking, Skiing}
U3 = {not France, Cuisine}
U4 = {Montreal, Canada, Reading}

Personalized
E1 = {Montreal, Canada}
E2 = {EJC}
E3 = {not France}
E4 = {Montreal, Canada}

Non-personalized
E = {Montreal, Canada, EJC, France}
Desired properties of an ad explanation

audience = (Montreal & Canada) or EJC or not France

U1 = {Montreal, Canada, Fishing, Driving}

U2 = {EJC, Hiking, Skiing}

U3 = {not France, Cuisine}

U4 = {Montreal, Canada, Reading}
Desired properties of an ad explanation

audience = (Montreal & Canada) or EJC or not France

U1 = {Montreal, Canada, Fishing, Driving}
U2 = {EJC, Hiking, Skiing}
U3 = {not France, Cuisine}
U4 = {Montreal, Canada, Reading}

Complete
E = {Montreal, Canada, EJC, France}

Succinct (incomplete, yet useful)
E = {Montreal}

Which attribute to show? based on prevalence? sensitivity?
Desired properties of an ad explanation

audience = (Montreal & Canada) or EJC or not France

U1 = {Montreal, Canada, Fishing, Driving}
U2 = {EJC, Hiking, Skiing}
U3 = {not France, Cuisine}
U4 = {Montreal, Canada, Reading}
Desired properties of an ad explanation

audience = (Montreal & Canada) or EJC or not France

U1 = {Montreal, Canada, Fishing, Driving}
U2 = {EJC, Hiking, Skiing}
U3 = {not France, Cuisine}
U4 = {Montreal, Canada, Reading}

Correct
E = {Montreal, Canada}
(vs. Incorrect)
E = {Montreal, Hiking}
Measurement methodology

- Build tool that collects ads and ad explanations on Facebook
  - Collection of real-world data
  - Controlled ad campaigns
Sample of explanations

<table>
<thead>
<tr>
<th>Category</th>
<th>Example of explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGE/GENDER/LOCATION</strong></td>
<td>You’re seeing this ad because MAHABIS wants to reach PEOPLE ages 24 AND OLDER who LIVE OR WERE RECENTLY IN FRANCE. This is information based on your Facebook profile and where you’ve connected to the internet.</td>
</tr>
<tr>
<td><strong>BEHAVIORS</strong></td>
<td>One of the reasons why you’re seeing this ad is because we think that you may be in the &quot;GMAIL USERS&quot; audience. This is based on what you do on Facebook.</td>
</tr>
<tr>
<td><strong>BIOGRAPHICAL DATA</strong></td>
<td>One reason why you’re seeing this ad is that BOREDOM THERAPY wants to reach people who SPEAK &quot;ENGLISH (US)&quot;. This is based on information from sources such as your Facebook profile.</td>
</tr>
<tr>
<td></td>
<td>One reason you’re seeing this ad is that AEGEAN AIRLINES wants to reach people with RELATIONSHIP STATUS &quot;ENGAGED&quot; on their Facebook profiles.</td>
</tr>
<tr>
<td></td>
<td>One reason why you’re seeing this ad is that EY CAREERS wants to reach people with THE SCHOOL/UNIVERSITY UNIVERSIT??T DES SAARLANDES - SAARLAND UNIVERSITY listed on their Facebook profiles.</td>
</tr>
<tr>
<td></td>
<td>One reason you’re seeing this ad is that ATENAO - TRANSLATION agency wants to reach people with THE EDUCATION LEVEL &quot;DOCTORATE DEGREE&quot; listed on their Facebook profiles.</td>
</tr>
<tr>
<td><strong>DATA BROKERS</strong></td>
<td>One reason you’re seeing this ad is that CANAL FRANCE wants to reach people who are part of an audience created based on data provided by ACXIOM. Facebook works with data providers to help businesses find the right audiences for their ads.</td>
</tr>
<tr>
<td><strong>DEMOGRAPHICS</strong></td>
<td>One of the reasons why you’re seeing this ad is because we think that you may be in the &quot;MILLENNIALS&quot; audience. This is based on what you do on Facebook.</td>
</tr>
<tr>
<td><strong>INTERESTS</strong></td>
<td>One reason why you’re seeing this ad is that ACER wants to reach people interested in ELECTRONIC MUSIC, based on activity such as liking Pages or clicking on ads.</td>
</tr>
<tr>
<td><strong>LIKED ADVERTISER’S PAGE</strong></td>
<td>One reason you’re seeing this ad is that SHAUN T wants to reach people who like their page.</td>
</tr>
<tr>
<td><strong>LOOKALIKE AUDIENCE</strong></td>
<td>One reason why you’re seeing this ad is that AUTODESK STUDENTS wants to reach people who may be similar to their customers.</td>
</tr>
<tr>
<td><strong>MOBILE DATA</strong></td>
<td>One reason why you’re seeing this ad is that CDU SAARBR??KEN-SCHEIDT wants to reach people WHO WERE RECENTLY NEAR THEIR BUSINESS. This is based on information from your Facebook profile and your mobile device.</td>
</tr>
<tr>
<td><strong>PII-BASED INFORMATION</strong></td>
<td>One reason you’re seeing this ad is that AAAS - THE AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE wants to reach people who have visited their website or used one of their apps. This is based on customer information provided by AAAS - THE AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE.</td>
</tr>
<tr>
<td></td>
<td>One reason you’re seeing this ad is that ACTIMEL added you to a list of people they want to reach on Facebook. They were able to reach you because you’re on their customer list or you’ve provided them with your contact information off of Facebook.</td>
</tr>
<tr>
<td></td>
<td>One reason you’re seeing this ad is that ABOUT YOU added you to an audience of people they want to reach on Facebook. This is based on activity such as watching their Facebook videos, sharing links to their website on Facebook and interacting with their Facebook content.</td>
</tr>
<tr>
<td><strong>SOCIAL NEIGHBORHOOD</strong></td>
<td>One reason why you’re seeing this ad is that CARTIER wants to reach people whose friends like their Page.</td>
</tr>
</tbody>
</table>
Evaluation of properties

- Personalized
- Incomplete (at most one attribute)
  - most prevalent attribute
  - demographics > interest > behaviors
- Sometime incorrect!
Evaluation of properties

- Personalized
- Incomplete (at most one attribute)
  - most prevalent attribute
    - demographics > interest > behaviors
- Sometime incorrect!

vulnerable to attackers!
Evaluation of properties

- Personalized
- Incomplete (at most one attribute)
  - most prevalent attribute
  - demographics > interest > behaviors
- Sometime incorrect!
Outline

- The ad targeting process
- The data inference process
Data inference process

Facebook actions (e.g., likes, clicks, posts)

Web browsing (online but outside Facebook)

Offline data

Ad platform

Inputs          Black-box          Outputs

Motherhood

New mover

Likely to engage in Politics (Liberal)
Explanations of the data inference process
Desired properties of a data explanation

- Specificity
- Completeness
- Correctness
Measurement methodology

• Build tool that collects the Ad Preference Page daily
  ➤ Collection of real-world data
  ➤ Controlled ad campaigns
Evaluation of properties

• Most explanations are vague

<table>
<thead>
<tr>
<th>Pattern</th>
<th>No explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td>You have this preference because you liked a Page related to <em>interest</em></td>
<td>2,866</td>
</tr>
<tr>
<td>You have this preference because you clicked on an ad related to <em>interest</em></td>
<td>2,095</td>
</tr>
<tr>
<td>You have this preference because we think it may be relevant to you based on what you do on Facebook, such as pages you’ve liked or ads you’ve clicked</td>
<td>785</td>
</tr>
<tr>
<td>You have this preference because you installed the app <em>app – name</em></td>
<td>107</td>
</tr>
</tbody>
</table>
| This is a preference you added.                                         | 2               

• Explanations are incomplete
AdAnalyst

http://adanalist.mpi-sws.org/

Disable AdBlockPlus on Facebook!
Have Facebook in english!
Discussions

• What else would you like to see in explanations?
• Can you think of other desired properties?
• What is the purpose of explanations for you?
Backup slides
Al Jazeera English

Website: https://www.facebook.com/7382473689
Advertiser Category: Company, Organization or Institution/ Media/News Company
Nb Likes: 10.51M
Nb of Ads I received: 1
I received the first ad on: 10/10/2017
I received the last ad on: 10/10/2017
The advertiser targeted me for: PII-based targeting,
The advertiser targeted other users for: PII-based targeting, CNN (Interests),
Great Big Story

Website:
https://www.facebook.com/1404914223144340

Advertiser Category: Company, Organization or Institution/ Media/News Company

Nb Likes: 4.761M

Nb of Ads I received: 1
I received the first ad on: 10/10/2017
I received the last ad on: 10/10/2017

The advertiser targeted me for: PII-based targeting,
The advertiser targeted other users for: Lookalike Audience, PII-based targeting, Acxiom (Data Brokers), Adventure travel (Interests), Airbnb (Interests), Asia (Interests), Documentary movies (Interests), Food Network (Interests), Frequent International Travelers (Behaviors), Frequent international travelers (Behaviors), Italy (Interests), The New York Times (Interests), Travel (Interests),
**Time:** 13/10/2017 10:09  
**Advertiser:** Western Union  
**Advertiser Category:** Product/Service  
**Explanation:** One reason you're seeing this ad is that Western Union wants to reach people who speak "Romanian". This is based on information from sources such as your Facebook profile. There may be other reasons you're seeing this ad, including that Western Union wants to reach people ages 18 and older who live or were recently in Germany. This is information based on your Facebook profile and where you've connected to the internet.  
**Compact explanation:** Romanian (Biographical Data)  
**Users also received this ad for:** Romanian (Biographical Data)  
**The advertiser targeted other users for:** Bollywood movies (Interests), Romanian (Biographical Data)  

---

**Time:** 24/09/2017 15:07  
**Advertiser:** Martin Schulz  
**Advertiser Category:** Politician  
**Explanation:** One reason you're seeing this ad is that Martin Schulz wants to reach people who speak "English (US)". This is based on information from sources such as your Facebook profile. There may be other reasons you're seeing this ad, including that Martin Schulz wants to reach people ages 18 and older who live near Saarbrücken, Saarland. This is information based on your Facebook profile and where you've connected to the internet.  
**Compact explanation:** English (US) (Biographical Data)  
**Users also received this ad for:** English (US) (Biographical Data)  
**The advertiser targeted other users for:** College (Interests), English (UK) (Biographical Data), English (US) (Biographical Data)  

---

Die neue Art des Bargeldsendens entdecken! westernunion.com Lassen Sie Ihre Freunde und Familie mehr Spaß haben – mit unseren verbessernten Wechselkurs...

<table>
<thead>
<tr>
<th>Parent</th>
<th>Reach</th>
<th>Sibling leaf categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>D/Politics (US)</td>
<td>179M</td>
<td>Likely To Engage in Politics (Conservative), Likely To Engage in Politics (Liberal), Likely To Engage in Politics (Moderate), US Politics (Conservative), US Politics (Liberal), US Politics (Moderate), US Politics (Very Conservative), US Politics (Very Liberal)</td>
</tr>
<tr>
<td>I/Family and relationships</td>
<td>138M</td>
<td>Dating, Family, Fatherhood, Friendship, Marriage, Motherhood, Parenting, Weddings</td>
</tr>
<tr>
<td>B/Consumer Classification/India</td>
<td>3100</td>
<td>(A) Affinity for High Value Goods/India, (A+B) Affinity for Mid-High Value Goods/India</td>
</tr>
<tr>
<td>D/Parents/All Parents</td>
<td>59M</td>
<td>(0-12 months) New Parents, (01-02 Years) Parents with Toddlers, (03-05 Years) Parents with Preschoolers, (06-08 Years) Parents with Early School Age Children, (08-12 Years) Parents with Preteens, (13-18 Years) Parents with Teenagers, (18-26 Years) Parents with Adult Children, Expectant parents, Parents (All)</td>
</tr>
<tr>
<td>B/Charitable donations</td>
<td>75M</td>
<td>Animal welfare, Arts and cultural, Environmental and wildlife, Health, Political</td>
</tr>
<tr>
<td>B/Financial/Spending methods</td>
<td>140M</td>
<td>1 Line of Credit, 2 Lines of Credit, 3, Active credit card user, Any card type, Bank cards, Gas, department and retail store cards, High-end department store cards, Premium credit cards, Primarily cash, Primarily credit cards, Travel and entertainment cards</td>
</tr>
<tr>
<td>B/Purchase behavior/Store types</td>
<td>34M</td>
<td>High-end retail, Low-end department store</td>
</tr>
<tr>
<td>B/Residential profiles</td>
<td>5.5M</td>
<td>Recent homebuyer, Recent mortgage borrower</td>
</tr>
<tr>
<td>B/Automotive/New vehicle buyers (Near market)/Style</td>
<td>102M</td>
<td>Crossover, Economy/compact, Full-size SUV, Full-size sedan, Hybrid/alternative fuel, Luxury SUV, Luxury sedan, Midsize car, Minivan, Pickup truck, Small/midsize SUV, Sports car/convertible</td>
</tr>
<tr>
<td>B/Purchase behavior/Health and beauty</td>
<td>90M</td>
<td>Allergy relief, Antiperspirants and deodorants, Cosmetics, Cough and cold relief, Fragrance, Hair care, Health and wellness buyers, Men’s grooming, Oral care, Over-the-counter medication, Pain relief, Skin care, Sun care, Vitamins</td>
</tr>
<tr>
<td>B/Automotive/Owners/Vehicle age</td>
<td>95M</td>
<td>0/1 year old, 1/15 years old, 16/20 years old, 2 years old, 3 years old, 4/5 years old, 6/10 years old, Over 20 years old</td>
</tr>
<tr>
<td>D/Home/Home Ownership</td>
<td>26M</td>
<td>First time homebuyer</td>
</tr>
<tr>
<td>B/Residential profiles</td>
<td>5.1M</td>
<td>New mover</td>
</tr>
<tr>
<td>B/Residential profiles</td>
<td>3M</td>
<td>Likely to move</td>
</tr>
<tr>
<td>B/Charitable donations</td>
<td>34M</td>
<td>All charitable donations, Cancer Causes, Children’s Causes, Veterans</td>
</tr>
<tr>
<td>D/Financial/Net Worth/Liquid assets (Acxiom)</td>
<td>74M</td>
<td>$1-$24,999, $100K-$249K, $1MM-$2MM, $25,000-$49,999, $250K-$499K, $2MM-$3MM, $3MM+, $50,000-$99,999, $500K-$1MM</td>
</tr>
<tr>
<td>B/Purchase behavior/Home and garden (DLX)</td>
<td>25M</td>
<td>Entertaining, Home renovation, Organization, Tools</td>
</tr>
<tr>
<td>B/Purchase behavior/Buyer profiles (Experian)</td>
<td>86M</td>
<td>Coupon users</td>
</tr>
<tr>
<td>B/Financial/Banking (Epsilon)</td>
<td>2.1M</td>
<td>Credit union member</td>
</tr>
</tbody>
</table>